GENERAL ETHICS

Killing people: what Kant could have said about suicide and euthanasia but did not

I Brassington

J Med Ethics 2006;**32**:571–574. doi: 10.1136/jme.2005.015420

An agent who takes his own life acts in violation of the moral law, according to Kant; suicide, and, by extension, assisted suicide are therefore wrong. By a similar argument, and with a few important exceptions, killing is wrong; implicitly, then, voluntary euthanasia is also wrong. Kant's conclusions are uncompelling and his argument in these matters is undermined on considering other areas of his thought. Kant, in forbidding suicide and euthanasia, is conflating respect for persons and respect for people, and assuming that, in killing a person (either oneself or another), we are thereby undermining personhood. But an argument along these lines is faulty according to Kant's own standards. There is no reason why Kantians have to accept that self-killing and euthanasia are contrary to the moral law. Even if some Kantians adhere to this doctrine, others can reject it.

Kant thinks that killing ourselves when life bodes ill is wrong. In itself, this does not mean that voluntary euthanasia is wrong, as euthanasia includes the intervention of another person. But if killing yourself is wrong, we may naturally think that killing other people (even for their own benefit or at their request) or getting others to kill you are things about which morality may have something to say. After all, it would appear to be the will to self-destruction that is problematic, and we may suppose that who executes this will makes little difference.

Kant's arguments for the wrongness of self-killing are, however (I claim), unconvincing, and the factors that undermine my confidence in Kant's arguments are Kantian themselves. But if Kant is unsuccessful in his attempt to show the wrongness of suicide, it becomes unclear why we should worry too much about enlisting someone else's help as a step towards suicide. Having said this, even if we can establish the permissibility of suicide, we will not necessarily have learnt anything about the permissibility of euthanasia thereby; by the same token, admitting the impermissibility of suicide gives no de facto indication that euthanasia is wrong, as the necessary participation of others in euthanasia allows a logical distinction to be drawn between euthanasia and suicide. Secondly, I want to trace what, if anything, the first part of the argument tells us about euthanasia.

The arguments that Kant[1] offers for the wrongness of suicide in the *Grounding* take two forms depending on whether we are arguing from the perspective of the first or second formulation of the categorical imperative (hereafter CI—additionally, I shall abbreviate the ''nth formulation of the categorical imperative'' to ''CIn''). These forms should give the same conclusion, as the various formulations of the categorical imperative indicate different ways of looking at the same basic rule.

Under the rubric of CI1, the rule against suicide is (allegedly) derived from the principle of universalisation. We should perhaps note at the outset that Kant's argument refers solely to suicide based on a desire to avoid evil, a term I take to be wide enough to include such evils as suffering; there is no reason why it should forbid the self-killing of a person who, guilt-ridden after committing a crime and agreeing with Kant's defences of judicial execution, but in a part of the world in which there is no court and no hangman, passes and executes a death sentence on himself: for the present purposes, I shall treat the word ''suicide'' as indicating ''self-killing to avoid suffering'' only.

Kant's claim is that a suicide indicates a maxim along the lines of ''From self-love I make as my principle to shorten my life when its continued duration threatens more evil than it promises satisfaction'',[1] but this is not something that lends itself to universalisation. After all, Kant reasons, nature places in us a feeling of self-love to stimulate the furtherance of life,[1] and the promotion of life through self-love can therefore be regarded as a law of nature. Meanwhile, to use self-love against life would necessitate the simultaneous promotion of death through self-love as a law of nature, and this would point to a contradiction in that law. Because a law—be it of nature or any other type—cannot contradict itself, we must suppose that the promotion of death through self-love violates the laws of nature and is therefore correctly called ''wrong''. (This, incidentally, is why parajudicial self-killing would be lawful: it is not carried out from a feeling of self-love.)

Is this compelling? I do not find it to be. For one thing, the wrongness in suicide lies in a contradiction in the laws of nature, not in the laws of nature per se; additionally, there is a difference between a contradiction ''in'', and a contradiction ''of'' laws that I shall exploit later. Now, to get rid of the wrongness implied by a contradiction in a law, all we have to do is resolve the contradiction; if we deny that the promotion of life through self-love is a law of nature, we can claim that suicide includes no violation, perversion or contradiction in any

Correspondence to:
I Brassington, School of Law, University of Manchester, Manchester, UK; Iain.Brassington@manchester.ac.uk

Received 2 December 2005
In revised form
10 January 2006
Accepted for publication
13 January 2006

natural law. And it seems clear that, regardless of whether a principle of self-love is motivationally ert or inert in any particular case, in no case does it promote my continued life. For sure, a native feeling of self-love may be sufficient to prevent me from killing myself; if, however, I do kill myself, this indicates nothing but that my native feeling of self-love is no longer sufficient. By contrast, Kant seems to be envisaging a picture in which agents may be debating the attractions of suicide, but may refrain from it because it dawns on them that they are blessed with self-love. I do not find this model sufficiently psychologically plausible to merit the effort of rejecting it.

So there does not seem to be a contradiction in any law of nature implied by a maxim of suicide. Nevertheless, perhaps Kant does himself a disservice with his account of self-love. We may be able to help him out if we can make a case for there being more than one kind of self-love and for these kinds of self-love being able to come into conflict. Availing ourselves of some of Rousseau's vocabulary (although not his use), we may be able to distinguish between ''amour propre''—which in this case I shall treat as a self-love based on presently occurrent desires to escape evil—and ''amour de soi''—which I shall treat as self-love based on genuine interests that promotes continued life. Here, it would seem that one kind of self-love may come into conflict with another; the supposedly single law of nature may mask two different and conflicting kinds of self-love. This would provide psychological plausibility; however, the price would be that a contradiction in the laws of nature is replaced by a conflict between two principles, and, however messy such a conflict may be, it is important that we keep hold of the idea that there would have to be a contradiction in a single law of nature if we are to establish wrongness via CI1. There is nothing worrisome or non-universalisable about a conflict of principles or motivational factors—a shopkeeper may be torn between his desire for wealth and his sense of honesty, for instance. The wrongness of dishonesty is not derived from the fact that honesty conflicts with his avarice, but from the fact that dishonesty is non-universalisable. The point is that, as suicide motivated by amour propre can be universalised, if it is wrong, it must be because there is something else going on.

The same point can be made in another way. If suicide motivated by self-love really did imply a serious contradiction in the laws of nature, notwithstanding that deriving wrongness from that would be naturalistically fallacious, it would be difficult to see how that suicide would be possible at all. Thus, considerations of whether to kill ourselves would turn out to be monkishly scholastic. The currency of impermissibility rests on possibility. If suicide is possible, then either it must be compatible with the laws of nature that determine our actions tout court—a possibility that Kant has denied, however problematically—or else we must be beholden to two kinds of law. And, of course, the idea that we are of dual nature and that moral laws determine the will and the laws of physics determine that which is unwilled about us is something that Kant[1] (passim, esp. Ak IV, 441ff) embraces. Indeed, the possibility of overcoming the laws of physical motion is crucial to Kant's thought: arguably, his whole output is haunted by the question of how morality is possible if we are simply lumps of warm fleshy material in the Newtonian world of experience (Kant,[2] p Bxxix.[3]).

Now, to the extent that Kant gives us reason to believe that we derive the laws of morality from that side of our nature that is non-physical, suicide implies nothing more than a contradiction of the laws of physics that determine the body by the (moral) laws that determine the will. As I argued a moment ago, however, a contradiction of one set of laws by another is no hindrance to Kantian morality. Indeed, without the means of escape from the physical world and the deterministic laws therein, there would be no way for an agent to be a moral agent; Kant's morality depends on us being at least partially non-physical. Again, the idea that the impermissibility of suicide can be established through a direct appeal to CI1 or the laws of nature seems to be wanting.

Further along the same track, a conceivable moral distinction exists between willing the destruction of our bodies and willing the destruction of our selves; even if the second cannot be universalised, I see no reason to suppose that the first cannot. People who commit suicide can, therefore, admit that the self will perish as a result of the destruction of the body that ails them, but could claim none the less (and on Kantian grounds) that perishing is no more than a foreseen but unintended consequence of their action. Kant himself hints at a denial that the self would perish anyway: ''the hypothesis of the spirituality of the rational beings of this world, according to which the body could remain dead on Earth and yet the same person still be living … is … congenial to reason,[and] not merely because it is impossible to conceive [of] matter endowed with thought … '';[4] this, I think, amounts to a confession that killing the body is not self-destructive. Thus, even if (arguendo) a desire to end one's life out of self-love indicates a contradiction in natural laws, and even if this is sufficient to justify the ascription of a predicate like wrong, we can still insist without leaving Kantian thought that a termination of embodiment out of self-love does not imply a will to end our lives or ourselves out of self-love. Overall, the argument against the permissibility of suicide that Kant builds around CI1 fails.

The argument that grows out of CI2 also fails. Here, Kant[1] wants to say that a suicide would be ''making use of his person merely as a means to maintain a tolerable condition until the end of his life''. This is incompatible with the moral law, which demands that we treat people as an end in themselves, and never wholly as a means. Personhood, Kant thinks, gives a creature a dignity that is beyond price; no matter what advantage we may gain by treating a person as a means to an end, this will never be sufficient to offset the value of their personhood. Quite why this should be is, I think, clearest when we view CI2 through the prism of CI1—just because CI1 cannot establish the impermissibility of suicide directly, it does not follow that it cannot do so indirectly. The argument would work like this: in treating people as a mere means, I am denying their personhood. But a maxim like ''deny personhood'' cannot be universalised without amounting to a denial of my own personhood—that is, my very capacity to form a maxim in the first place. So to treat people as mere means violates the moral law—and this naturally extends to treating myself as a thing.

Again, there is no need to deny Kantian premises to argue that none of this will yield a duty to reject suicide. For the structure of Kantian moral thought, and the move that Kant[1] is forced to make for morality to be possible, means that we have to think of ourselves as belonging at the same time to the sensible world and to a radically distinct intelligible world. It is my super—sensibility that grants me my status as a source of value, gives me the capacity for autonomy and so on; it is this status that puts me beyond price.

Yet, Kant has a problem here. For my knowledge of myself as an individual agent and of my personhood is sensible—a matter of mere experience; my apparent knowledge of myself is knowledge of just another appearance in a world of appearances (Kant,[2] p A38/B55). But the personhood that puts me beyond price has its roots in the intelligible world, and has, for that reason, no appearance. In fact, in itself, it need not even be thought of as individuated: personhood

may be something in me without being contained within me. Hence, I need not restrict respect for the personhood within a person to respect for personhood as it appears in that person; the personhood is transcendent. But if personhood transcends the individual, whereas killing a person may indicate a refusal to respect personhood, it does not have to. Now, if I can coherently believe that the personhood within me belongs to sensibility at the same time that I believe that the personhood in me transcends sensibility—and I can—I can easily separate the destruction of myself from the destruction of the personhood in the person that I am. As it is the personhood in the person that I am that demands respect, suicide need not be thought of as violating the moral laws that demand respect for personhood. I can, in other words, claim that willing the disposal of my (sensible or experienced) existence in no way implies disposing of the invaluable part of me. Suitably extended, this sort of argument would also do away with Kant's[5 6] claims in his *Lectures on ethics* and in the *Metaphysics of morals* that suicide amounts to the willed disposal of personhood. It does not.

Of course, the thought that my experience of personhood may bear no relationship to the reality of the personhood that is in me may not come easily. But there is nothing incoherent about a separation of my experience of personhood from its reality—in Kant's words, ''I can *think* whatever I please, provided only that I do not contradict myself, that is, provided my concept is a possible thought''(Kant,[2] p Bxxvi n)—and so such a distinction can, without intellectual impropriety, influence my moral decision making. Moreover, the interpretation of Kant on which this argument relies is not particularly outré: the idea that the personhood that we know in ourselves is nothing more than the appearance of a noumenon that need not or cannot be limited to ourselves is close to the reading of Kant that provided Schopenhauer[7 8] with the foundation to his claim that the distinction between people—indeed, between a person and the rest of the universe—was merely a matter of illusion.

In a sense, such a line of argument seems to clear space for a restatement of the classical Kantian refusal of suicide's permissibility; for if a serious distinction needs to be drawn between my personhood and the noumenal personhood in me, then the very legitimacy of thinking about this priceless characteristic as ''mine'' in the first place is rendered questionable: that, after all, would subsume the noumenal to the phenomenal. On such a reading, one thing that we may want to say about the destruction of the personhood in me through suicide is that it is not mine to destroy, and we may want to tell a story about how that makes suicide wrong. But, of course, adopting such an argument would equally well lead to the thought that the apparent destruction of the self is, in fact, nothing more than the destruction of the apparent self: the apparent destruction of the self is no such thing anyway. For Schopenhauerian pessimists, this means that not even suicide provides a reliable escape from their rather bleak lives; transcendental idealism provides no reason to commit suicide. For more traditional Kantians and post-Kantians, however, the same sort of reasoning seriously weakens the reasons that we might have not to.

The point is that, in killing the person that I am, I have set that individual life at a price. But in doing so, I am not necessarily making any dent in the integrity, dignity or value of the personhood therein, because I can claim that my experience of my individual personhood is importantly separable from the non-individuated reality thereof, and it is the noumenal, non-individuated personhood in an individual person that makes the creature that he is worthy of respect. As the personhood can be thought of as transcending the necessarily phenomenal, sensible person, I can coherently claim that I am still setting personhood above price, even though I may be bringing about the death of a particular person.

Once again, Kant's denial of the permissibility of suicide is found wanting; this leaves the way clear for an argument for the permissibility of suicide on the basis of appeals to autonomy (which is also founded in our intelligible sides). There need be no clash between respect for autonomy and absolute inner worth,[9] and it would, in fairness, be surprising if there were a clash—the two are, after all, ways of making sense of the same noumenal characteristic. The implications for assisted suicide should be clear: if suicide is not wrong, then there would seem to be no reason why soliciting help for suicide should be wrong. Equally, if suicide is permissible, I think we may suppose that it is also permissible to respond to a call for assistance by assisting. Although there is no reason to suppose that it may be obligatory actually to help someone die, neither is there any basis for a refusal to do so in the wrongness of a proposed suicide, because there is no wrongness.

As I mentioned earlier, a moral difference needs to be drawn between suicide or assisting suicide and killing a person; even if I provide the means to commit suicide, I have not killed you. So we may ask what the position is in Kantian thought on euthanasia. Classically, the response has been that, as suicide is impermissible, so is euthanasia. Notwithstanding what has come before, I think that this move is sometimes made too hastily; but, even so, if suicide is permissible, we are still not able to say with certainty that euthanasia is. For one thing, my reason for carrying out euthanasia may accord with your morally permissible request to die, but may at the same time derive from my desire to get my hands on your possessions. For another, making an appeal to a desired way out of suffering will not establish euthanasia's permissibility, as the rightness of an action does not hinge on its goodness: Kant claims (admittedly, in another context, but one that translates well enough) that ''it is monstrous to suppose that we can have a right to do wrong in the direst physical distress''[10] and (again, in a different context) tries to show with his example of the murderer at the door that the goodness of fulfilling our desire to prevent a murder by lying must, morally speaking, be subjugated to the duty of truthfulness.[11] Having said this, the wrongness of euthanasia cannot lie simply in the fact that it is killing a person, as there are times when this is permitted—nay, required—by Kant: the unlawful killing of another must, he thinks, be punished by death;[12] it is this sort of consideration that leads me to suspect overhastiness in leaping from statements about suicide to statements about euthanasia.

If killing is sometimes permissible, it cannot contravene CI1; whereas we may not desire a world with universalised homicide, such a world would be coherent. If and when killing is wrong, it must be because of something that it signifies. I believe that the problem with euthanasia, if there is one, is best explained by CI2: it is that, in carrying out euthanasia, we are treating a person wholly as a means to an end: we have desired that we want to minimise suffering, and we have chosen to kill as a means to that.

We should note, though, that even if we are inclined to accept this argument against euthanasia, it still only works if our action is motivated by something such as a desire to reduce the net level of suffering in the world and if the euthanasia and the person euthanised are thereby treated as a means to that end; it only rules out euthanasia that is involuntary or non-voluntary. (Non-voluntary euthanasia here covers cases in which a person may want to be killed but this fact plays no part in our decision to kill him.) In those cases when a person wants us to kill him, when that desire is necessary and sufficient to motivate us to kill him, and in

which we have no other motivationally ert desire to kill him, it would be difficult to sustain the charge that we are acting in such a way so as to make a person a means to our end. And so it seems that voluntary euthanasia can be fairly easily shown to be compatible with CI2. Nor do I see any reason to worry about problems with reconciling this with membership and sovereignty in the Kingdom of Ends—and as all the formulations of the CI are supposed to be practically equivalent, this is how it should be.

I am certainly not trying to suggest that Kant would allow us to do anything to a person as long as that person had autonomously requested it. If a person tried to sell himself to me as a slave, Kantian thought would still give me a reason to refuse the transaction, for it is in the nature of slavery that it fails to respect the personhood in a person. But, in acceding to a request for euthanasia, it is possible that I am treating a person as a means to an end and thereby undermining the personhood in him—I would certainly be treating a person as a means to an end if I tried to convince him that he had a duty to be killed—but it is not necessarily the case that this is what I am up to. Buying you as a slave is, at least sometimes, worse than killing you.

One potential glitch remains: given that moral laws are supposed to be obtained universally, might not we say properly that morality is agent indifferent and that, for that reason, it does not matter who performs an action? If we do want to say this, we may be led to the thought that a person who kills himself as a way of ending his own suffering is acting in the same way as a person who kills another person as a way of ending that other's suffering—that is, that one may treat oneself as an other and, wrongly, instrumentally. But in this case, just because we are (allegedly) dual, we could interpret this situation as one in which the will directs the body to behave in a specified way to achieve a particular end; this is not problematic in any other case—it is not wrong for my will to treat my body as a means to the satisfaction derived from, say, eating ice cream—and so it is unclear why it should be a problem in this one. As I have argued, because we are (allegedly) dual, there is no reason why a will that wills the destruction of the body must be willing its own destruction. My willing my own physical destruction for your benefit may imply that I have thingified myself, and this may well contravene the CI. But willing my physical destruction to escape suffering seems to do no such thing.

The conclusion I am drawn towards is this: the ''official'', mainstream interpretation of Kant, according to which he forbids all suicides, all assisted suicide and all euthanasia, is simply not tenable, and in rejecting it, we do not even have to step outside of Kantianism. We can reasonably easily out-Kant Kant.

Competing interests: None.

REFERENCES

1 **Kant I**. *Grounding for the metaphysics of morals*. Indianapolis: Hackett, 1993, Ak IV, 422.
2 **Kant I**. *Critique of pure reason*. Basingstoke: Palgrave Macmillan, 2003.
3 **Gardner S**. *Kant and the critique of pure reason*. London: Routledge, 2003:16–8.
4 **Kant I**. *Religion within the boundaries of mere reason*. Cambridge: Cambridge University Press, 2004:128n.
5 **Kant I**. *Lectures on ethics*. Whitstable: Methuen, 1979:149.
6 **Kant I**. The doctrine of virtue. In: Kant I, eds. *The metaphysics of morals*. Cambridge: Cambridge University Press, 1998:422–3.
7 **Schopenhauer A**. *The world as will and representation*, Vol 1. New York: Dover, 1969 (passim)).
8 **Schopenhauer A**. *On the basis of morality*. Indianapolis: Hackett, 1995, passim).
9 **Häyry M**. The tension between self governance and absolute inner worth in Kant's moral philosophy. *J Med Ethics* 2005;**31**:645–7.
10 **Kant I**. On the common saying: ''This may be true in theory, but it does not apply in practice''. In: Kant I, eds. *Political writings*. Cambridge: Cambridge University Press, 2004:81.
11 **Kant I**. On the supposed right to lie from philanthropic concerns. In: Kant I, eds. *Grounding for the metaphysics of morals*. Indianapolis: Hackett, 1993:426.
12 **Kant I**. The doctrine of right. In: Kant I, eds. *The metaphysics of morals*. Cambridge: Cambridge University Press, 1998:337.